Information and Computation 210 (2012) 31–39

Contents lists available at SciVerse ScienceDirect

Information and Computation

www.elsevier.com/locate/yinco

Sparse reliable graph backbones

Shiri Chechik[a], Yuval Emek[b], Boaz Patt-Shamir[c,*,1], David Peleg[a,2]

[a] Department of Computer Science and Applied Mathematics, Weizmann Institute of Science, Rehovot 76100, Israel
[b] Computer Engineering and Networks Laboratory, ETH Zurich, CH-8092 Zurich, Switzerland
[c] School of Electrical Engineering, Tel Aviv University, Tel Aviv 69978, Israel

ARTICLE INFO

Article history:
Received 24 November 2010
Revised 3 October 2011
Available online 26 October 2011

Keywords:
Network reliability
Sparse subgraphs
Tutte polynomial

ABSTRACT

Given a connected graph G and a failure probability $p(e)$ for each edge e in G, the *reliability* of G is the probability that G remains connected when each edge e is removed independently with probability $p(e)$. In this paper it is shown that every n-vertex graph contains a sparse *backbone*, i.e., a spanning subgraph with $O(n \log n)$ edges whose reliability is at least $(1 - n^{-\Omega(1)})$ times that of G. Moreover, for any pair of vertices s, t in G, the (s,t)-*reliability* of the backbone, namely, the probability that s and t remain connected, is also at least $(1 - n^{-\Omega(1)})$ times that of G. Our proof is based on a polynomial time randomized algorithm for constructing the backbone. In addition, it is shown that the constructed backbone has nearly the same *Tutte polynomial* as the original graph (in the quarter-plane $x \geqslant 1$, $y > 1$), and hence the graph and its backbone share many additional features encoded by the Tutte polynomial.

1. Introduction

Finding a sparse subgraph that approximately preserves some key attribute of the original graph is fundamental to network algorithms: any lazy network manager would find the capability to maintain fewer links in a large network a precious gift. This can also be considered from the perspective of identifying a set of redundant edges in a graph. Whether an edge is redundant or not depends of course on the attributes that should be preserved. Spanners [15,16] for example, approximately preserve pairwise distances in graphs, with a trade-off spectrum between the quality of approximation and the number of edges in the spanner. The general graph attribute we focus on in the current paper is connectivity under random edge failures.

Specifically, we consider the classical setting of *network reliability*, defined over a graph G whose edges e are associated with *failure* probabilities $p(e)$. The *reliability* of G is the probability that G remains connected when each edge e of G is removed independently with probability $p(e)$. Clearly, the reliability of a graph is monotone non-increasing with respect to edge removal. We seek a sparse spanning subgraph (containing all vertices and only a small subset of the edges) of G, referred to henceforth as a *backbone*, whose reliability is almost as good as that of G.

Our main result is a randomized algorithm for constructing a backbone with $O(n \log n)$ edges that approximates the reliability of G to within a (multiplicative) factor of $1 - n^{-\Omega(1)}$, where n denotes the number of vertices. The randomized algorithm allows edge multiplicities, so the original graph G may have significantly more than $\binom{n}{2}$ edges. This construction is tight: we show that there are graphs whose reliability cannot be approximated to within any positive factor by any subgraph with significantly less than $n \log n$ edges. Moreover, the backbone graph approximates not only the *all-terminal*

* Corresponding author.
E-mail address: boaz@eng.tau.ac.il (B. Patt-Shamir).

[1] Supported in part by Israel Science Foundation (grant 1372/09) and by the Israel Ministry of Science and Technology.
[2] Supported in part by the Israel Ministry of Science and Technology.

0890-5401/$ – see front matter © 2011 Elsevier Inc. All rights reserved.
doi:10.1016/j.ic.2011.10.007

variant of the reliability (the probability that the whole graph remains connected), but also the (s,t)-*reliability* of G for any two vertices s and t, defined as the probability that s and t remain in the same connected component. Our construction is presented first for the *homogeneous* case, where the failure probability of every edge is some constant $0 < p < 1$, and then extended to the general *heterogeneous* case, assuming that there aren't "too many" edges whose failure probabilities are very close to 1 (see Section 3.2 for a precise statement).

It turns out that our backbone also provides a good approximation for the *Tutte polynomial*.[3] Specifically, in the quarter-plane $x \geqslant 1$, $y > 1$ the Tutte polynomial of the backbone approximates the Tutte polynomial of the original graph to within a factor of $1 \pm n^{-\Omega(1)}$ after multiplying by a (trivially calculated) normalizing factor that accounts for the different number of edges. Since the Tutte polynomial encodes many interesting features of the graph (including its reliability), this result seems to indicate that our backbone construction provides a good representation of the graph in some deeper sense.

Related work. Network reliability is a fundamental problem in operations research since the early days of that discipline [13]; see the survey [2] for a comprehensive account. It is also well known in the area of computational complexity; various versions of the network reliability problem are listed among the 14 basic #P-complete problems[4] presented in [19]. In particular, both the all-terminal reliability problem and the (s,t)-reliability problem are known to be #P-hard even when the failure probabilities $p(e)$ are homogeneous. [10] establishes a fully polynomial time randomized approximation scheme (FPRAS) for the problem of evaluating the probability that the graph disconnects under random edge failures. Although this disconnection probability is simply one minus the reliability of the graph, the algorithm of [10] does not translate to a (multiplicative) approximation for the problem of evaluating the reliability. In fact, the approximability of the all-terminal reliability and the (s,t)-reliability problems is still an open question.

A notion somewhat related to ours is that of graph *sparsifiers* [17,18]: An n vertex weighted graph H is said to be a κ-*sparsifier* of an n vertex weighted graph G if $x^T L_G x \leqslant x^T L_H x \leqslant \kappa \cdot x^T L_G x$ for every vector $x \in \mathbb{R}^n$, where L_H and L_G are the Laplacian matrices of H and G, respectively. Sparsifiers are a generalization of the *compressed graphs* of [4], that approximately preserve the total weight of edges crossing any cut in the original graph. Indeed, the graph compression condition corresponds to the sparsifier condition restricted to vectors $x \in \{0,1\}^n$.

One is interested in constructing sparse sparsifiers (hence the name) and the state of the art in that context is the recent construction of $(1+\epsilon)$-sparsifiers with $O(n/\epsilon^2)$ edges presented in [3]. Note that unlike the backbone constructed in the current paper, sparsifiers are not required to be subgraphs of the original graph. Furthermore, even if a sparsifier edge is present in the original graph, its weight may be different. In fact, there exist unweighted graphs for which every good sparsifier must introduce edges of widely varying weights [18].

A brief overview of the Tutte polynomial is given in Section 5. Here we comment that the computational complexity of evaluating the Tutte polynomial on various points $(x,y) \in \mathbb{R}^2$ is almost completely understood. The problem admits an efficient algorithm if $(x,y) \in \{(1,1),(-1,-1),(0,1),(-1,0)\}$ or if $(x-1)(y-1) = 1$; otherwise it is #P-hard [8]. An FPRAS exists for the $y > 0$ portion of the "Ising" hyperbola $(x-1)(y-1) = 2$ [9]; and unless RP = NP, an FPRAS does not exist if $x < -1$ or if $y < -1$ except for the aforementioned easy-to-compute points, the ray $x < -1$, $y = 1$, and the $y < -1$ portion of the hyperbola $(x-1)(y-1) = 2$ [7]. An FPRAS also exists for the quarter-plane $x \geqslant 1$, $y \geqslant 1$ if the minimum degree in G is $\Omega(n)$ [1] and for the half-plane $y > 1$ if the size of a minimum cut in G is $\Omega(\log n)$ [10].

Technique. Our backbone construction samples each edge with probability inverse proportional to its *strength*, a parameter closely related to edge connectivity. This technique was introduced in [4] for the construction of compressed graphs. In [4], the weights of the selected edges are then modified to meet the graph compression condition. This cannot be done when constructing a backbone: we can only remove edges, and are not allowed to change intrinsic attributes (namely failure probabilities) of the remaining ones. Nevertheless, we show that with high probability, the resulting backbone approximately preserves the reliability of the original graph. The main ingredient in our analysis is the fact that graphs with logarithmic edge connectivity are highly reliable [12,10]. (Note that we do not make any assumptions on the connectivity of the original graph.) The Tutte polynomial analysis is slightly more involved and it essentially relies on an observation of [1] combined with a theorem of [10].

Paper organization. The remainder of this paper is organized as follows. Section 2 includes the preliminaries used throughout the paper. The backbone construction is presented in Section 3 and the matching lower bound is established in Section 4. In Section 5 we prove that our backbone also provides a good approximation for the Tutte polynomial.

2. Preliminaries

Unless stated otherwise, all graphs mentioned in this paper are undirected and not necessarily simple (i.e., they may contain parallel edges and self-loops). We denote the vertex set and edge set of a graph G by $V(G)$ and $E(G)$, respectively. The graph *induced* on G by a vertex subset $U \subseteq V(G)$ is $G(U) = (U, E(G) \cap (U \times U))$. The graph *induced* on G by an

[3] The Tutte polynomial $T_G(x,y)$ is a bivariate polynomial whose coefficients are determined by the graph G. See Section 5 for details.

[4] The complexity class #P consists of the counting problems whose decision versions are in NP.